

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2001, 18:06:14 ; Search time 2667.65 Seconds
(without alignments)
4145.761 Million cell updates/sec

Title: US-09-784-340-3_COPY_197_911
Perfect score: 715
Sequence: 1 atgaggtctgcaagtcagc.....gtttatagtaagcattag 715

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba1: *
2: gb_ba2: *
3: gb_ba3: *
4: gb_in1: *
5: gb_in2: *
6: gb_in3: *
7: gb_om: *
8: gb_ov: *
9: gb_pat1: *
10: gb_pat2: *
11: gb_ph: *
12: gb_pi1: *
13: gb_pi2: *
14: gb_pi3: *
15: gb_pi4: *
16: em_ba1: *
17: em_ba2: *
18: em_fun: *
19: em_htgo_hum: *
20: em_htgo_inv: *
21: em_htgo_rod: *
22: em_htg_hum1: *
23: em_htg_hum2: *
24: em_htg_hum3: *
25: em_htg_hum4: *
26: em_htg_hum5: *
27: em_htg_hum6: *
28: em_htg_hum7: *
29: em_htg_hum8: *
30: em_htg_inv1: *
31: em_htg_inv2: *
32: em_htg_other: *
33: em_htg_rod: *
34: em_hum1: *
35: em_hum2: *
36: em_hum3: *
37: em_hum4: *
38: em_hum5: *
39: em_hum6: *
40: em_hum7: *
41: em_in: *
42: em_om: *
43: em_or: *

44: em_ov: *
45: em_pat: *
46: em_ph: *
47: em_pi: *
48: em_rod: *
49: em_sts: *
50: em_sy: *
51: em_un: *
52: em_v1: *
53: gb_sts1: *
54: gb_sts2: *
55: gb_sts3: *
56: gb_sy: *
57: gb_un: *
58: gb_v12: *
59: gb_v12: *
60: gb_htg1: *
61: gb_htg2: *
62: gb_htg3: *
63: gb_htg4: *
64: gb_htg5: *
65: gb_htg6: *
66: gb_htg7: *
67: gb_htg8: *
68: gb_htg9: *
69: gb_htg10: *
70: gb_htg11: *
71: gb_htg12: *
72: gb_htg13: *
73: gb_htg14: *
74: gb_htg15: *
75: gb_htg16: *
76: gb_htg17: *
77: gb_htg18: *
78: gb_htg19: *
79: gb_htg20: *
80: gb_htg21: *
81: gb_htg22: *
82: gb_htg23: *
83: gb_htg24: *
84: gb_htg25: *
85: gb_pi1: *
86: gb_pi2: *
87: gb_pi3: *
88: gb_pi4: *
89: gb_pi5: *
90: gb_pi6: *
91: gb_pi7: *
92: gb_pi8: *
93: gb_pi9: *
94: gb_pi10: *
95: gb_pi11: *
96: gb_pi12: *
97: gb_pi13: *
98: em_ba3: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	703.8	98.4	212904	66	AC021146	AC021146 Homo sapi
C 2	539.8	75.5	2786	89	AK025587	AK025587 Homo sapi
C 3	407.8	57.0	139015	69	AC025488	AC025488 Homo sapi
C 4	407.8	57.0	169246	62	AC011254	AC011254 Homo sapi
C 5	406	56.8	138232	63	AC012582	AC012582 Homo sapi
C 6	391.6	54.8	226077	63	AC013296	AC013296 Homo sapi
C 7	348.8	48.8	2547	94	AF175221	AF175221 Cavia por
C 8	190.2	26.6	2090	97	HSU08854	HSU08854 Human UDP g

```

Sequencing vector: M13; 88%
Sequencing vector: plasmid; 12%
Chemistry: dye-primer ET; 88% of reads
Chemistry: dye-terminator Big Dye; 12% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 199729 bases at least Q40
Consensus quality: 203731 bases at least Q30
Consensus quality: 206540 bases at least Q20
Insert size: 213000; agarose-fp
Insert size: 210604; sum-of-contigs
Quality coverage: 3.75 in Q20 bases; sum-of-contigs
Quality coverage: 3.83 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1
1565: 1564: contig of 1564 bp in length
1565: gap of unknown length
1665: 3085: contig of 1421 bp in length
3086: gap of unknown length
3186: 5134: contig of 1949 bp in length
5135: gap of unknown length
5136: 8169: contig of 2835 bp in length
8170: gap of unknown length
8270: 10341: contig of 2072 bp in length
10342: gap of unknown length
13615: 13614: contig of 3173 bp in length
13616: gap of unknown length
13715: 17089: contig of 3375 bp in length
17189: gap of unknown length
17190: 20701: contig of 3512 bp in length
20702: 20801: gap of unknown length
20802: 25001: contig of 4200 bp in length

```

*	25012	29120	gap of unknown length
*	29021	29120	contig of 4236 bp in length
*	29121	33356	contig of 4236 bp in length
*	33357	33456	gap of unknown length
*	33457	36902	contig of 3446 bp in length
*	36903	37002	gap of unknown length
*	37003	42975	contig of 5973 bp in length
*	42976	43075	gap of unknown length
*	43076	48338	contig of 5264 bp in length
*	48340	48439	gap of unknown length
*	48440	56529	contig of 8090 bp in length
*	56530	56629	gap of unknown length
*	56630	65213	contig of 8584 bp in length
*	65214	65313	gap of unknown length
*	74715	74715	contig of 9402 bp in length
*	74716	74815	gap of unknown length
*	74816	88546	contig of 13731 bp in length
*	88547	103367	gap of unknown length
*	103368	103467	contig of 14721 bp in length
*	103468	118167	contig of 14700 bp in length
*	118168	118267	gap of unknown length
*	118268	132765	contig of 14498 bp in length
*	132766	132865	gap of unknown length
*	132866	154092	contig of 21227 bp in length
*	154093	154192	gap of unknown length
*	154193	182144	contig of 27952 bp in length

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH068N14
----- Summary Statistics -----

```


Oy	525	aggaagaataggaaggaagcctctgggaaacttcacgcccacacttccatgactctgagc	584
Db	361	AGGCATATATGAGGAAGACTGTGGAAACTTCACAGCTCACTTTCCTATCTACTGTGCC	420
Oy	585	tatgacagagactaacagacagacgaatgacacctctctgaaagagtaaaaaatcaatgcttc	644
Db	421	TATGACAGAGACTTAACAGACAGATGACCTTCTTGAAAGATGATAAAATTCATAGCTTTC	480
Oy	645	agtttctgccaacttcctggatccaaggattacgaatacatcatllttgggaaagaatttataag	704
Db	481	AGTTTGTGTCACCTCTGGATTCAGGATTCAGACTATCACTTTTGGGAAAGATTATAG	540
Oy	705	taagcattag 715	
Db	541	TAAAGCATTAG 551	

[illegible]

REFERENCE	TITLE	JOURNAL	AUTHORS
1 (bases 1 to 139015)	Homo sapiens chromosome 4, clone RP11-790I12	Unpublished	Birren,B., Linton,L., Nussbaum,C., Lander,E.
2 (bases 1 to 139015)			Birren,B., Linton,L., Nussbaum,C., Lander,E., Abraham,H., Allen,N.

TITLE Direct Submission
JOURNAL Submitted (09-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: MIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence.submissions@genome.wi.mit.edu

 Project Information
 Center project name: L6943
 Center clone name: 790.I.12

 Summary Statistics
 Sequencing vector: M13; M77815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads

```
Assembly program: Pirap; version 0.960731
Consensus quality: 132515 bases at least Q40
Consensus quality: 135966 bases at least Q30
Consensus quality: 137953 bases at least Q20
Insert size: 139000; agarose-fp
Insert size: 138015; sum-of-contigs
Quality coverage: 4.4 in Q20 bases; agarose-fp
Quality coverage: 4.4 in Q20 bases; sum-of-contigs
```

NOTE: This is a working draft sequence. It currently consists of 11 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

*	1	1340	1339:	contig of 1339 bp in length
*		1440	1439:	gap of 100 bp
*		4292	4291:	contig of 2852 bp in length
*		4392	4391:	gap of 100 bp
*		7033	7032:	contig of 2641 bp in length
*		7133	7132:	gap of 100 bp
*		13335	13344:	contig of 6202 bp in length
*		13435	13434:	gap of 100 bp
*		23885	23884:	contig of 10450 bp in length
*		23985	23984:	gap of 100 bp
*		36107	36106:	contig of 12122 bp in length
*		36207	36206:	gap of 100 bp
*		51820	51819:	contig of 15613 bp in length
*		51920	51919:	gap of 100 bp
*		65522	65521:	contig of 13502 bp in length
*		65532	65531:	gap of 100 bp
*		82012	82011:	contig of 16490 bp in length
*		82112	82111:	gap of 100 bp
*		98266	98265:	contig of 16154 bp in length
*		98366	98365:	gap of 100 bp
*		139015	139015:	contig of 40650 bp in length

FEATURES
source

```

misc_feature      /clone.lib="RRC1-11 Human Male BAC"
                  1..1339
misc_feature      /note="assembly_fragment"
                  1440..4291
misc_feature      /note="assembly_fragment"
                  4392..7032
misc_feature      /note="assembly_fragment"
                  7133..13334
misc_feature      /note="assembly_fragment"
                  clone.end:T7
                  vector_side:right"
misc_feature      13435..23884
                  /note="assembly_fragment"
misc_feature      23985..36106
                  /note="assembly_fragment"
misc_feature      36207..51819
                  /note="assembly_fragment"
misc_feature      51920..65421
                  /note="assembly_fragment"
                  clone.end:SP6
                  vector_side:left"
misc_feature      65522..82011
                  /note="assembly_fragment"
misc_feature      82112..98265
                  /note="assembly_fragment"
misc_feature      98366..139015
                  /note="assembly_fragment"

```


* NOTE: This is a working draft sequence. It currently
* consists of 53 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	1640:	contlg of 1640 bp	in length
1641	1740:	gap of 100 bp	in length
1741	2943:	contlg of 1203 bp	in length
2944	3043:	gap of 100 bp	in length
3044	4997:	contlg of 1854 bp	in length
4998	4997:	gap of 100 bp	in length
4998	6114:	contlg of 1117 bp	in length
6115	6214:	gap of 100 bp	in length
6215	7665:	contlg of 1451 bp	in length
7666	7765:	gap of 100 bp	in length
7766	9188:	contlg of 1423 bp	in length
9189	9288:	gap of 100 bp	in length
9289	11371:	contlg of 2083 bp	in length
11372	11471:	gap of 100 bp	in length
11472	13194:	contlg of 1723 bp	in length
13195	13294:	gap of 100 bp	in length
13295	14424:	contlg of 1030 bp	in length
14425	14424:	gap of 100 bp	in length
14425	15511:	contlg of 1087 bp	in length
15512	15611:	gap of 100 bp	in length
15612	17562:	contlg of 1751 bp	in length
17563	17462:	gap of 100 bp	in length
17463	19182:	contlg of 1720 bp	in length
19183	19282:	gap of 100 bp	in length
19283	20717:	contlg of 1435 bp	in length
20718	20817:	gap of 100 bp	in length
20818	22066:	contlg of 1249 bp	in length
22067	22166:	gap of 100 bp	in length
22167	23790:	contlg of 1624 bp	in length
23791	23890:	gap of 100 bp	in length
23891	25462:	contlg of 1572 bp	in length
25463	25562:	gap of 100 bp	in length
25563	27562:	contlg of 2000 bp	in length
27563	27662:	gap of 100 bp	in length
27663	29314:	contlg of 1652 bp	in length
29315	29414:	gap of 100 bp	in length
29415	31317:	contlg of 1903 bp	in length
31318	31417:	gap of 100 bp	in length
31418	33429:	contlg of 2012 bp	in length
33430	33529:	gap of 100 bp	in length
33530	35133:	contlg of 1604 bp	in length
35134	35233:	gap of 100 bp	in length
35234	37156:	contlg of 1923 bp	in length
37157	37256:	gap of 100 bp	in length
37257	39661:	contlg of 2405 bp	in length
39662	39761:	gap of 100 bp	in length
39762	41496:	contlg of 1735 bp	in length
41497	41596:	gap of 100 bp	in length
41597	43309:	contlg of 1713 bp	in length
43310	43409:	gap of 100 bp	in length
43410	45445:	contlg of 2036 bp	in length
45446	45545:	gap of 100 bp	in length
45546	47118:	contlg of 1573 bp	in length
47119	47218:	gap of 100 bp	in length
47219	49435:	contlg of 2217 bp	in length
49436	49535:	gap of 100 bp	in length
49536	52546:	contlg of 3011 bp	in length
52547	52646:	gap of 100 bp	in length
52647	53529:	contlg of 883 bp	in length
53530	53629:	gap of 100 bp	in length
53630	55789:	contlg of 2160 bp	in length
55790	55889:	gap of 100 bp	in length
55890	59029:	contlg of 3140 bp	in length
59030	59129:	gap of 100 bp	in length

*	59130	61713: contig of 2584 bp	in length
*	61714	61813: gap of 100 bp	
*	61814	63696: contig of 1883 bp	in length
*	63697	63796: gap of 100 bp	
*	63797	65790: contig of 1994 bp	in length
*	65791	65890: gap of 100 bp	
*	65891	67845: contig of 1955 bp	in length
*	67846	67945: gap of 100 bp	
*	67946	69933: contig of 1988 bp	in length
*	69934	70033: gap of 100 bp	
*	70034	72807: contig of 2774 bp	in length
*	72808	72907: gap of 100 bp	
*	72908	75815: contig of 2908 bp	in length
*	75816	75915: gap of 100 bp	
*	75916	79940: gap of 4025 bp	in length
*	79941	80040: gap of 100 bp	
*	80041	83640: contig of 3600 bp	in length
*	83641	83740: gap of 100 bp	
*	83741	86921: contig of 3181 bp	in length
*	86922	87021: gap of 100 bp	
*	87022	91044: contig of 4023 bp	in length
*	91045	91144: gap of 100 bp	
*	91145	94259: contig of 3115 bp	in length
*	94260	94359: gap of 100 bp	
*	94360	96910: contig of 2251 bp	in length
*	96911	97010: gap of 100 bp	
*	97011	100209: contig of 3199 bp	in length
*	100210	100309: gap of 100 bp	
*	100310	105003: contig of 4694 bp	in length
*	105004	105103: gap of 100 bp	
*	105104	110031: contig of 4228 bp	in length
*	110032	110131: gap of 100 bp	
*	110132	115352: contig of 5221 bp	in length
*	115353	115452: gap of 100 bp	
*	115453	120961: contig of 5509 bp	in length
*	120962	121061: gap of 100 bp	
*	121062	126155: contig of 5094 bp	in length
*	126156	126255: gap of 100 bp	
*	126256	130511: contig of 4256 bp	in length
*	130512	130611: gap of 100 bp	
*	130612	138232: contig of 7621 bp	in length
FEATURES			
source	Location/Qualifiers		
	1..138232		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone_lib="RPI1-3C24"		
	1..1640		
misc_feature	/note="assembly_fragment"		
	141..2943		
misc_feature	/note="assembly_fragment"		
	3044..4897		
misc_feature	/note="assembly_fragment"		
	4998..6114		
misc_feature	/note="assembly_fragment"		
	6215..7665		
misc_feature	/note="assembly_fragment"		
	7766..9188		
misc_feature	/note="assembly_fragment"		
	9289..11371		
misc_feature	/note="assembly_fragment"		
	11472..13194		
misc_feature	/note="assembly_fragment"		
	13295..14324		
misc_feature	/note="assembly_fragment"		
	14425..15111		
misc_feature	/note="assembly_fragment"		
	15613..17362		
misc_feature	/note="assembly_fragment"		
	17463..19182		
misc_feature	/note="assembly_fragment"		
	19283..20717		
misc_feature	/note="assembly_fragment"		

FEATURES

Source

misc_feature	1.1640
misc_feature	/note="assembly_fragment"
misc_feature	1741.2943
misc_feature	/note="assembly_fragment"
misc_feature	3044.4897
misc_feature	/note="assembly_fragment"
misc_feature	4998.6114
misc_feature	/note="assembly_fragment"
misc_feature	6215.7665
misc_feature	/note="assembly_fragment"
misc_feature	7766.9188
misc_feature	/note="assembly_fragment"
misc_feature	9289.11371
misc_feature	/note="assembly_fragment"
misc_feature	11472.13194
misc_feature	/note="assembly_fragment"
misc_feature	13295.14324
misc_feature	/note="assembly_fragment"
misc_feature	14425.11511
misc_feature	/note="assembly_fragment"
misc_feature	15612.17362
misc_feature	/note="assembly_fragment"
misc_feature	17463.19182
misc_feature	/note="assembly_fragment"
misc_feature	19283.20717
misc_feature	/note="assembly_fragment"

misc_feature

```

misc_feature      20818..22066
                  /note="assembly_fragment"
misc_feature      22167..23790
                  /note="assembly_fragment"
misc_feature      23891..25462
                  /note="assembly_fragment"
misc_feature      25563..27562
                  /note="assembly_fragment"
misc_feature      27663..29314
                  /note="assembly_fragment"

Query Match
Best Local Similarity 56.8%; Score 406; DB 63; Length 138232;
Matches 542; Conservative 0; Mismatches 155; Indels 13; Gaps 3;

QY 1 atgagatctgaacagtcagcttgatgtatctgcctccagcagctcttcgtgtgctgt 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 101021 ATGAGTCTGAGAGAGGCTGTTTACAAATTCCTACCTCAGCTCTGCTGCTGCTGT 100962

QY 61 ggattcttgaggaaagctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100961 GGATTCTTGAGAGAGGCTGTTTACAAATTCCTACCTCAGCTCTGCTGCTGCTGT 100902

QY 121 gtcattctgaagagctcatgtatgtatgtatgtatgtatgtatgtatgtatgtatgt 179
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100901 AATATTTTGGAGAACTCATAGAGAGGGGTCAACAGTAACACTGTTGACTCACACAC 100842

QY 180 gccctgttaattgactacaggaagcccttcctgcattgaaatttgaggtgtccatgcc 239
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100841 AGCCTTTTCAATGACTAGAGGAGAGCCCTTCTGCATTCGAACTTGTAGCTGCTGCTGCC 100782

QY 240 acagagcaga-----acagaaagaaatgaatatattgttgactagctgtgagatgt 290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100781 ACATGCAAGAGTGAATTTTGACAAATGATGATGAAATTTCTGGACCTGCTGAAATGT 100722

QY 291 ctgtcagagcttaacacactgtgcaatcagctgtataaataaagattttgttgtaaat 350
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100721 CTGTGCAAGCTTGACCTGCTGCGGAGTCAAGAAAGAACTGAATATTTT---CTGAAATC 100665

QY 351 aagaggaactttaaaatgatgtgtgagagcttattctacacagacactatgtgaaga 410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100664 TAGAGAAATTTTAAACCTCTGTGTAGAGTTTGGCTTCAATCAGACACTCTTTGAAGAT 100605

QY 411 gctcaaggaaacacactaagatgataagccttataagacccgtgatacccggtgtgagacct 470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100604 GCAAGAGGAACCCCACTAAACTGTAGCTTATAGACTCTGCCATTCCTGAGGAGGTT 100545

QY 471 gatgctgagctgtctccagctccctttgtgtcacactagaactctctaaagagcaga 530
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100544 TATTGCTGAGTTCCTTGCAATTCCTTTTGTGTCAAACTTAGAGCTTCTGTAGCTGGCAA 100485

QY 531 tatgagcgaagctgtgtggaacttccagctccacttctctatgtacagctgtgctatgac 590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100484 TATTGAGTAGAGCTGTGGGAAACTTCCACCTCCACTTCTATGTATCTGTGCTTANAGAC 100425

QY 591 aggaactaagcagaagatgaccttctgtgaaagagtaaaaaatcaatgcttcagttt 650
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100424 AGGACTTAACGGAATTAAGACTTCTTCTGGAAAAAGTAAAAATTAACAATGCTGTGCATCTT 100365

QY 651 gtctcactgtgattcaagattacagactatcttttgggaagagttt 700
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100364 TTTTAACCTTGCTGCCTCCAGATTTTAACTTTTGTGATTTTGGAACATTTT 100315

RESULT 6
AC013296/c AC013296 226077 bp DNA HTG 13-JUL-2000
LOCUS Homo sapiens clone Rpl1-3M18, LOW-PASS SEQUENCE SAMPLING.
DEFINITION AC013296
ACCESSION AC013296.5 GI:9121251
VERSION HTG: HTGS_PHASE0.
KEYWORDS human.
SOURCE ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 226077)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone Rpl1-3M18
Unpublished
2 (bases 1 to 226077)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,A., Beckery,L., Boguslavsky,L., Boukhalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collamore,A.,
Cooke,P., DeRellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Headford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kahn,L., Karatas,A., Klein,J.,
Lewoczky,J., Lieu,C., Locke,K., Macdonald,P., Margulis,N.,
McEwan,P., McQuirk,A., McKernan,K., McLaughlin,J., Melidori,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollard,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (06-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6514003.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
Project Information
Center project name: L2775
Center clone name: 3_M_18

* NOTE: This record contains 246 individual
* sequencing reads that have been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1
815 914: gap of 100 bp
915 1704: contig of 790 bp in length
1705 1804: gap of 100 bp
1805 2602: contig of 798 bp in length
2603 2702: gap of 100 bp
2703 3484: contig of 782 bp in length
3485 3584: gap of 100 bp
3585 4378: contig of 794 bp in length
4379 4478: gap of 100 bp
4479 5257: contig of 779 bp in length
5258 5357: gap of 100 bp
5358 6150: contig of 793 bp in length
6151 6250: gap of 100 bp
6251 7038: contig of 788 bp in length
7039 7138: gap of 100 bp
7139 7940: contig of 802 bp in length
7941 8040: gap of 100 bp
8041 8869: contig of 829 bp in length
8870 8969: gap of 100 bp
8970 9790: contig of 821 bp in length
9791 9890: gap of 100 bp
9891 10695: contig of 805 bp in length
10696 10795: gap of 100 bp
10796 11592: contig of 797 bp in length

```


[illegible]

REFERENCE	1 (bases 1 to 2090)
AUTHORS	Green,M.D., Olturu,E.M. and Tepfly,T.R.
TITLE	Stable expression of a human liver UDP-glucuronosyltransferase (UGT2B15) with activity toward steroid and xenobiotic substrates
JOURNAL	Drug Metab. Dispos. 22 (5), 799-805 (1994)
MEDLINE	95136867
REFERENCE	2 (bases 1 to 2090)
AUTHORS	Green,M.D.
TITLE	Direct Submission
JOURNAL	Submitted (13-APR-1994) Mitchell D. Green, Department of Pharmacology, The University of Iowa, Iowa City, IA 52242, USA
FEATURES	Location/Qualifiers
source	1..2090
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="HeBa"
	/clone_1ib="Human Liver Unizap cDNA library from Stratagene"
	/tissue_type="liver"
sig_peptide	22..90
	/gene="UGT2B15"
CDS	22..1614
	/gene="UGT2B15"
	/EC_number="2.4.1.17"
	/codon_start=1
	/product="UDP glucuronosyltransferase precursor"
	/protein_id="AAC50077.1"
	/db_xref="GI:475759"
	/translation="MSLWMTSVFLIOLSCYFSSGCGKVLVMPREXSHNMKTILEELVRQHEVTVLTSSASTLVAKSSAIAKEVYPPTSLKNLEDSLILKIIDRWIGYSKNFMWSYEOSLOELCWERYDYXSNLKCDAYINIKMLMKLOESKDVIADANCGELLAELEFNIPFLISLFPSVGYTFEKKGGEPLEPSTVPVAVMSLSOOMEMETKNTIMHLPEFWFOIIDLKRWIDPFYSVLGRPTPLEETMGAAMWLIRYWDIEFPPLPNVDVFGLGHCPAKPAKPKEMEFEVSOGSENGIVFSGSMISNSESANMIASALOIQKVLMRFDAKKNPLTGSRITRYKMLPDNDLGHPKTAFTIHGNGNGYEAIYHGIPMWGIPLFAOHNDIIAHMKAGKAISVIDIRMTSSRLDNLNKSYINDPYVKENMYKISRTHHDQMCKLRDRAVFWIEFWFERMRKGAKLHVAAHNLTWLYOHSUDVIALACVAIVFTITRKCLCEFRKIATGKKRRD"
gene	22..1614
	/gene="UGT2B15"
mat_peptide	91..1611
	/gene="UGT2B15"
	/EC_number="2.4.1.17"
	/product="UDP glucuronosyltransferase"
BASE COUNT	649 a 384 c 435 g 622 t
ORIGIN	
Query Match	26.6%; Score 190.2; DB 97; Length 2090;
Best Local Similarity	56.0%; Pident No. 4.9e-38;
Matches 407; Conservative 0; Mismatches 308; Indels 12; Gaps 2;	
Oy	1 atgaagctgcagaatcagcttggtatttcctgcccgcagacctcttcgtg---gtggc 57
Dbb	19 AGGATGTCTCGAATATGGACGTCACACTTTCTTGCATACAGCTCAGTTGTACTTAGC 78
Oy	58 tqtgatctctcgaaagagctccctggctgagcccttgacacagagccattggtctaatac 117
Dbb	79 TCTGGAACCTGTGGAAAAGCTGCTATGTGGCCACAGAAATACAGCCATTGGATAATATG 138
Oy	118 aagtcatacttagaagaagcatatagttagaaggccaatagagtaaacagatatgaactaca 177
Dbb	139 AAGACAATCTGTGGAAGACCTGTGTCAAGGGGCTCATGAGTGACAGTGTGACATCTTCG 198
Oy	178 aagctctgtaaatgagcatagcaagaagctctgcatgaaattggaagtgtgcatactg 237
Dbb	199 GCTTCTACTCTTGTCAAAGCCAGTAATCATCTGCTATTAAATTAAGAAGTTATCTTACA 258
Oy	238 cc-----acagagacagaacagaagaataatgaaatatttgttaccatagctctgaa 288
Dbb	259 TCTTTAACTAAAAATGATTTTGGAGAATTTCTCTTCGAAAATTTTCGATAGATGATATAT 318
Oy	289 gtcttgcagagctatacaactctgacalcagttataaaatcaatgattttttgttgtaa 348

[illegible]

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P. Y., Zhao, B., Frengen, E., Taitano, M., Caranese, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bapac.med.buffalo.edu>)

VECTOR: pBACe3.6
NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is CTD-2005D20. Actual start of this clone is at base position 1 of RPI1-185H6; actual end is at base position 181117 of RPI1-185H6.

FEATURES

source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RPI1-185H6"
/clone_lib="RPCI-11"
1..71
repeat_region /rpt_family="ERV1"
271..444
repeat_region /rpt_family="ERV1"
472..592
repeat_region /rpt_family="ERV1"
593..938
repeat_region /rpt_family="MaLR"
939..1217
repeat_region /rpt_family="ERV1"
1617..3178
repeat_region /rpt_family="L1"
3189..5532
repeat_region /rpt_family="L1"
5533..5818
repeat_region /rpt_family="Alu"
5819..6208
repeat_region /rpt_family="L1"
6219..6642
repeat_region /rpt_family="ERV1"
6660..7156
repeat_region /rpt_family="ERV1"
9336..10141
repeat_region /rpt_family="ERV1"
10149..10219
repeat_region /rpt_family="L1"
10242..10508
repeat_region /rpt_family="ERV1"
10509..10978
repeat_region /rpt_family="L1"
10979..11006
repeat_region /rpt_family="ERV1"

```

```

repeat_region 10988..12442
/rpt_family="ERV1"
repeat_region 12446..12879
/rpt_family="ERV1"
repeat_region 12921..13150
/rpt_family="L1"
repeat_region 13187..13499
/rpt_family="L1"
repeat_region 13549..13976
/rpt_family="L1"
repeat_region 14032..14404
/rpt_family="L1"
repeat_region 14499..14598
/rpt_family="ERV1"
repeat_region 14628..16122
/rpt_family="L1"
repeat_region 16135..16537
/rpt_family="ERV1"
repeat_region 16615..16996
/rpt_family="L1"
repeat_region 16998..17293
/rpt_family="Alu"
repeat_region 17710..17766
/rpt_family="Alu"
repeat_region 17801..18334
/rpt_family="L1"
repeat_region 18343..18622
/rpt_family="L1"
repeat_region 18623..19027
/rpt_family="MaLR"
repeat_region 19028..20370
/rpt_family="L1"
repeat_region 20386..20698
/rpt_family="ERV1"
repeat_region 20699..20938
/rpt_family="ERV1"
repeat_region 20977..21022
/rpt_family="ERV1"
repeat_region 21043..22020
/rpt_family="ERV1"
repeat_region 22022..22314
/rpt_family="ERV1"
repeat_region 22321..22476
/rpt_family="L1"
repeat_region 22500..23925
/rpt_family="L1"
repeat_region 24128..24222
/rpt_family="ERV1"
repeat_region 24452..24625
/rpt_family="ERV1"
repeat_region 24645..24899
/rpt_family="L1"
repeat_region 24916..25128
/rpt_family="MaLR"
repeat_region 25129..25436
/rpt_family="Alu"
repeat_region 25437..25499
/rpt_family="MaLR"
repeat_region 27081..27433
/rpt_family="MaLR"
repeat_region 27437..28027
/rpt_family="L1"
repeat_region 28021..28967
/rpt_family="L1"
repeat_region 28954..29769
/rpt_family="L1"
repeat_region 29770..30074
/rpt_family="Alu"
repeat_region 30075..30629
/rpt_family="L1"
repeat_region 30631..31554
/rpt_family="L1"
repeat_region 31675..32162

```

```

repeat_region      /rpt_family="ERV1"
                    /rpt_family="Alu"
repeat_region      /rpt_family="ERV1"
                    /rpt_family="ERV1"
repeat_region      /rpt_family="L1"
                    /rpt_family="L1"
repeat_region      /rpt_family="L1"
                    /rpt_family="L1"
repeat_region      /rpt_family="L1"
                    /rpt_family="L1"
repeat_region      /rpt_family="L1"
                    /rpt_family="L1"
repeat_region      /rpt_family="L1"
                    /rpt_family="L1"
repeat_region      /rpt_family="L1"
                    /rpt_family="L1"

```

Query Match 26.6% Score 190.2; DB 87; Length 181117;
 Best Local Similarity 56.0% Pred. No. 7.3e-38;
 Matches 407; Conservative 0; Mismatches 308; Indels 12; Gaps 2;

```

QY 1 atgaggtctgacagtcagcttggatcttctgctcctgagctctctgt---gttgc 57
DB 86499 AGGATGTCCTGTAATGACGTCAGCTCTTCTGCTGATACAGCTGTTGTTACTTAA 86558
QY 58 tctgagctctgctggaagaatctcctgctgctgctgacatgagccatgcttaatgc 117
DB 86559 TCTGGAAGCTGTGGAAGGCTGCTAGTGTGCGCCACAGATACAGCCATTGATTAATG 86618
QY 118 aaggtatctcagaagagctatgtagagagccatgagtglaacaglatgactca 177
DB 86619 AAGACAATCTCTGGAAGAGCTGTTGTCAGAGGGGTCATGAGGTGATGACATCTTCG 86678
QY 178 aagcctctgtaattgactacaggaagcctctgctgactgaaattgagtgccatg 237
DB 86679 GCTTCTACTCTGTCACATGCTGTAATGCTGCTATTAATTAAGTTATTCCTACA 86738
QY 238 cc-----acagagcagaagaagaatgaaatattgttgactagctcgaat 288
DB 86739 TCTTAACTAAATGATTTGGAAAGTTCTCTTCTGAAATTTCTCATAGATGATATAT 86798
QY 289 gcttcgaggtctcaacctggaacatgataataaataaagattttgttgaa 348
DB 86799 GGTGTTTCAAAAATACATTTTGGTCATATTTTTCACAAATTACAGAATTTGTGGAA 86858
QY 349 ataaagagacttaaaatgtagtgtagagcttatctacacatcagacacttaag 408
DB 86859 TATTATGACTACAGTACAGGCTGTGTAAAGATGACAGTTTGAATAAGAACTTATGATG 86918
QY 409 aagctacaggaacaacaactacagatgaacgtatagaccctgtagtcccgctgagac 468
DB 86919 AAACGTACAGAGTCAAGTTTATGATGATCTTGCGAGATGCCCTTAACTCCCTGTGAG 86978
QY 469 ctgaggtcgaattgctcagctcccttctgctcactgaactgaactctcgaagagc 528
DB 86979 CTACTGCGTGAATATTAAATACATCCCTTCTGTCACAGTCTTGATTCCTGTTGGCTAC 87038
QY 529 aatatgagcgaagcgtgtaggaactccaagctccaacttccatgactgagcctag 568
DB 87039 ACATTTGGAAGAAGAGTGTGAGAGTCTGTCCTCTCTCATATGACCTGTTGTTATG 87098
QY 589 acagagctacagacagaatgaccccttcggaagaaglaaaaaatccaatgacttaagt 648
DB 87099 TCAGATTTAAGTGAATGATTTTCATGAGAGATGAAAAATATGATACATAGCTTT 87158
QY 649 tctgctcactctgagctcagaagtaacatcatttgggaagagcttaataag 708
DB 87159 TATTTTGACTTTTGTTCATAATTATGATCTGAGAGAGTGGACAGCTTTTATAGTGA 87218
QY 709 gcaatg 715
DB 87219 GTTCTAG 87225

```

```

RESULT 10
AF180322
LOCUS
DEFINITION
Homo sapiens UDP-glucuronosyltransferase 2B15 (UGT2B15) mRNA.
ACCESSION
AF180322
VERSION
AF180322
KEYWORDS
SOURCE
ORGANISM
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
Levesque, E., Beaulieu, M., Green, M.D.,ephly, T.R., Belanger, A. and
Hum, D.W.
TITLE
Isolation and characterization of UGT2B15(Y85): a
UDP-glucuronosyltransferase encoded by a polymorphic gene
JOURNAL
Pharmacogenetics 7 (4), 317-325 (1997)
MEDLINE
97439504
REFERENCE
2 (bases 1 to 2114)
AUTHORS
Levesque, E., Beaulieu, M., Green, M.D., ephly, T.R., Belanger, A. and
Hum, D.W.
TITLE
Direct Submission
JOURNAL
Submitted (24-AUG-1999) Molecular Endocrinology, CHUL Research
Center, 2705 Laurier Blvd, Ste-Foy, Quebec G1V 4G2, Canada
FEATURES
source
1..2114
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4q13"
1..2114
/gene="UGT2B15"
/allele="Y85"
39..1631
/gene="UGT2B15"
/codon_start=1
/product="UDP-glucuronosyltransferase 2B15"
/protein_id="A0D55093.1"
/db_xref="GI:3681246"
/translation="MSLRKTSVFLDIQ...VFSSGSGKVLVMPTEYSHWIMNKILE
ELVQGHVTVLTSASTLVNASKSA...EYPSYSLKNTLESLKILDMRYGVS
KNTFMSYFSQLCEIWEYDYDYNKRLCK...NKRKLMLKLOESKFPVILADLNPCEL
LAELNIPFVLSRVSQVTEPKNGGG...PSYVPMVMSLSDQMIFMERIKNIHM
LYFDWROIYIDKKDOYSEYLGPTI...EMKAEMLIRTYNDFEPRRLPNVD
FVGLHCKPAKPLKEMEFVOSSENG...VFSLSMSNSSEANIASALQIPQ
KYLWREDKRPMTLGSNRLKMLPND...SHPKKATITGGTNGIYEALYHGPVY
GIPLFADODHNAIHKAKGALSVDIR...SHDLNLALKSYINDVYEVNWKLSRIH
HDQPMKPLDRAVFMIEFVMBRKHGAKH...AHNLTWIQYHSLDVIAPLACVATVIFI
ITKPLCPFRKLAKTGKKKKND"
BASE COUNT
670 a 386 c 435 g 623 t
ORIGIN
Query Match 26.4% Score 188.6; DB 39; Length 2114;
Best Local Similarity 55.8% Pred. No. 1.3e-37;
Matches 406; Conservative 0; Mismatches 33; Indels 12; Gaps 2;
QY 1 atgaggtctgacagtcagcttggatcttctgctcctgctcctctctgt---gttgc 57
DB 36 AGGATGTCCTGTAATGACGTCAGCTCTTCTGCTGATACAGCTGTTGTTACTTAA 95
QY 58 tctgagctctgctggaagaatctcctgctgctgctgacatgagccatgacttaatgc 117
DB 96 TCTGGAAGCTGTGGAAGGCTGCTAGTGTGCGCCACAGATACAGCCATTGATTAATG 155
QY 118 aaggtatctcagaagagctatgtagagagccatgagtglaacaglatgactca 177
DB 156 AAGACAATCTCTGGAAGAGCTGTTGTCAGAGGGGTCATGAGGTGATGACATCTTCG 215
QY 178 aagcctctgtaattgactacaggaagcctctgctgactgaaattgagtgccatg 237
DB 216 GCTTCTACTCTGTCACATGCTGTAATGCTGCTATTAATTAAGAGTTATTCCTACA 275

```

```

QY 238 cc-----acagacagacagagaagaataatattgttgcactgctcgaat 288
Db 276 TCTTTAACTAAATAATTTTGGAGATCTCTTGGAAATTCGTAGATGATATAT 335
QY 289 gcttcgacaggtctcaactggaacatcagttataataatgaatttttctgaa 348
Db 336 GGTGTTTCAAAAATACATTTTGGTCATATTTTTCACAAATTAACAAGAAATGTGTGGGA 395
QY 349 ataagaggaacttaaaaatgattgtgtgagagcttattcctacacagacacttaaga 408
Db 396 TATTATGACTACAGTACAGTCTGTAAAGTGCAGTTTGAATTAAGAACTTATGATG 455
QY 409 aagctacagagaacccaactacagatgtaacgcttatalagacctgtgattcccgtagagac 468
Db 456 AACTCTACAGAGTCAAGAGTTTGCATTCATTCGACAGTCCCTTAATCCCTGTGTAG 515
QY 469 ctgagagctgagttgctccagtccttctgtgtcctacacacttagaactctcctaagagac 528
Db 516 CTACTGGCTGAACATATTTAAACATACCCCTTGTGTACAGTCTTGATGATCTCTGTGGCTAC 575
QY 529 aataagagcagaagctgtggaacactccagctccacttccatgtaactgtgacctag 588
Db 576 ACATTTGAGAAAGATGCTGAGAGATTTCTGTCCTCCATGCTGATGCTGTATG 635
QY 589 acagagcacaacagacagaatgaccttctggaagaagtaaaaaatcaatgcttcaagt 648
Db 636 TCAGAAATTAAGTCAATGATTTTCATGAGAGATTAATAAATATGATACATATGCTT 695
QY 649 ttgttcacactcgtgattcagattacgactatcatttctggaagaagtttaagtaag 708
Db 696 TATTGACTTTTGTTCATTAATTTATGATCTGAAGAAGAGTGGACCAAGTTTATATGAA 755
QY 709 gcaatag 715
Db 756 GTTCTAG 762

RESULT 11
AF081793 2091 bp mRNA PRI 18-AUG-1998
LOCUS Homo sapiens UDP glucuronosyltransferase 2B4 precursor (UGT2B4)
DEFINITION mRNA, UGT2B4*1L09, L396 variant allele, complete cds.
ACCESSION AF081793
VERSION AF081793.1 GI:3426331
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2091)
AUTHORS Jin, C.J., Miners, J.O., Lillywhite, K.J. and Mackenzie, P.I.
TITLE cDNA cloning and expression of two new members of the human liver
UDP-glucuronosyltransferase 2B subfamily
JOURNAL Biochem. Biophys. Res. Commun. 194 (1), 496-503 (1993)
MEDLINE 93326164
JOURNAL 2 (bases 1 to 2091)
REFERENCE Mackenzie, P.I.
AUTHORS Direct Submission
TITLE Submitted (30-JUL-1998) Clinical pharmacology, Flinders University
JOURNAL of South Australia, Bedford Park, SA 5042, Australia
FEATURES
source
1..2091
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
1..2091
/gene="UGT2B4"
/note="formerly UGT2B11"
/allele="UGT2B4*L109, L396 variant"
36..1622
/gene="UGT2B4"
/codon_start=1

```

```

/product="UDP glucuronosyltransferase 2B4 precursor"
/protein_id="A0C32272.1"
/db_xref="GI:3426332"
/translation="MSKMTSALLLLQLSCYSSGCGKVLWPEFSHMMNKTILD
ELVORGEHVLAASSASISFDPSPSTLKEVLPVSTRTFEDDILKOLYKRAEPLK
DTEWYSLSQVOEIMMTFENDILRRCFDIVSNKLMKRLDSRDVVLADVPFGEL
AELKIPFYSLRSPGYAIEKSGGILPSPVPVPMVMSOUMFIEKVMKIVL
VPEFNOIFPMKKMDQFSEVLRPTLLSTAKADIMLRNRYMDFPHPLPNEV
VGGLHCKPAPLPLKMKKEFVOSSENGVYVSELSVMSERRAVIASLAKIPQK
VLMFQDNKPDITGNATRIKNTKPTPNDLITGPKTRAFITGANGAYEAIHGIPWG
VPLADPDNIAMKKAAGAAVSDIFHTMSDTDLNALKYINDPLKENMKLSRIHH
DQYKPLDRAVFMIEFVMRHKGAKHLVNAHDLTFWYHSLDVTGELLCVATVIFII
TKCLFCVWKEFRTGKGRD"
BASE COUNT 638 a 401 c 437 g 615 t
ORIGIN
Query Match 26.3%; Score 188.4; D1: 88; Length 2091;
Best Local Similarity 55.8%; Pred. No. 1.4e-37;
Matches 404; Conservative 0; Mismatches 3.1; Indels 9; Gaps 2;
1 atgagctgacaagtcagcttggatttgcctccgcagctctgtg---gtggc 57
Db 33 AGGATGCTATGAATAATGAGACTTCAAGCTCTCTGCTGATACACAGAGCTGTACTTAC 92
QY 58 tggattctgtggaagaagctcgtgtgcccgtggaacatgagcattgactaagtc 117
Db 93 TCTGGAGAGTGTGGAAGAGTGCTGTGGCCACAGAAATTAAGCAGTGAATATA 152
QY 118 aagtcattctagaagagctcagtagaagagcattgaggaagattatgactaacca 177
Db 153 AAGACAATCCGTGAGAACCTTGTCCAGAGAGGTGAGAGTACGTGATGTGGACTTTCA 212
QY 178 aagccttgtaattgacctacagagaagcctctgcatgaatctgagtgagttccatag 237
Db 213 GCTTCATTTCTTTCGATGCCAACGCCCATCTACTCTTAATTTGAAGTTATCCTGTA 272
QY 238 ccagagacagacagagaagaatgaatattgttgcac---agctctga-----atgtc 291
Db 273 TCTTTAACTAAATAATGAGTTCATGAGTATTTATCAAGACAG: GGTAAAGAGATGGCAGAA 332
QY 292 ttgcagagcttcaactcggcgaatcagttataataatgaatttttctgtaata 351
Db 333 CTTCACAAAAGACACATTTTGGTCATATCTTTCACAGTACAGAAATCACTGTGACATTT 392
QY 352 agaggaacttaaaaatgattgtgtgagagcttattcctacacacagacacttaagaag 411
Db 393 AATGACTACTTAGAAGTCTCTTAAGATATAGTTTCAATAAGAACTGTATGAAGAA 452
QY 412 ctacagagaacccaactacagatgtaacgcttataagacctgtgattcccgtagagacctg 471
Db 453 CTACAGAGTCAAGATTTGATGTTGTTTTCGACAGCGCTGTTTCCCTTGTGTGAGCTG 512
QY 472 atgctgagttgttcacgtcccttctgtgtcacaacttagaacctctctaagagcaat 531
Db 513 CTGGCGGAGTACTTAATAAATACCTTTGTCATACAGCTCTCTCTCTCTCTGCTAGCA 572
QY 532 atgagcagaagctgtgtggaacttccagctccacttccatgtaactgtgcctatgaca 591
Db 573 ATTGAAAGAGATGTGAGAGACTTGTCCTCCCTCATATGCTGTGTATGATGA 632
QY 592 ggaactacagacagaatgaccttctggaagaagtaaaaaatcaatgcttcaagtttg 651
Db 633 GAACTAAGTACCAAAATGACTTTCATAGAGAGGTAAATAATATGATGATGCTTTAT 692
QY 652 ttccacttcgtgattcagagttacgactacatttctggaagaagatttaagtaagca 711
Db 693 TTTGAATTTGTTCCAAATATTTGACATGACAGAAAGTGGATCAGTTCTTCACTGAAGTT 752
QY 712 ttgag 715
Db 753 CTAG 756

```

```

RESULT 12
AF112112 2108 bp mRNA PRI 11-APR-1999
LOCUS Macaca fascicularis UDP-glucuronosyltransferase 2B19 precursor,
DEFINITION mRNA, complete cds.
ACCESSION AF112112
VERSION AF112112.1 GI:4580601
KEYWORDS crab-eating macaque.
SOURCE Macaca fascicularis.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 2108)
AUTHORS Belanger, G., Barbier, O., Hum, D. W. and Belanger, A.
TITLE Molecular cloning, expression and characterization of a monkey
steroid UDP-glucuronosyltransferase, UGT2B19, that conjugates
testosterone
JOURNAL Eur. J. Biochem. 260 (3), 701-708 (1999)
MEDLINE 99203465
REFERENCE 2 (bases 1 to 2108)
AUTHORS Belanger, G., Barbier, O., Hum, D. W. and Belanger, A.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-1998) Molecular Endocrinology, CHUL Research
Center, 2705 Laurier Blvd, Sainte Foy, Quebec G1V 4G2, Canada
FEATURES
Source Location/Qualifiers
1..2108
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
26..1612
/EC_number="2.4.1.17"
/note="UGT2B19"
/codon_start=1
/product="UDP-glucuronosyltransferase 2B19 precursor"
/db_xref="GI:4580602"
/translation="MSMKWTSAALLLIQLSCYLSFGCGKVLVMPFEFSHMNIKTILD
ELVORGHEVTLAVASTSLDPDNPSPLEKFEICPTSLTFEEODSVTOLVKRMSDRK
DPEWFLHVOEMMTYGMIRKCKDVSNNKLKKDOESRFDVLAIAISPCELL
AELKIPVYSLRFSFGTALKEHGGFELPSTVPVTKMLDQMTFMRVNMITYM
YDFEWQVVDVKNWDOFSKVLGRPTTLEIMAKELIMLRNWFQFPHLPNVEF
VGLHCKPAKPLPKMEEFVQSSDNGVVFSLGSVSNMSEERANVLSALAKIPQK
VLMRDGNKPDITGLNTOLIKWLPONDILGHKPTRAFTHGANGAYEALVIGIPWG
VLPADDPNINAHMKAKGAVALDPTMTSTLLNAIKTVINDPIYKEAMKLSITH
DQVPRDLRAVFWIEFVNRHAKKHLRVAAHDLTFQYHSLDYIGFLACVATVIFII
TGLFCVMKFFVRKKGRKD"
BASE COUNT 649 a 393 c 438 g 628 t
ORIGIN
Query Match 26.3%; Score 188.4; DB 89; Length 2108;
Best Local Similarity 55.8%; Pred. No. 1.4e-37;
Matches 404; Conservative 0; Mismatches 311; Indels 9; Gaps 2;
QY 1 atgaagctgcacaagtcagcctgtgtatcttcgtctcgtcgaagctcttcgtt--gtttgc 57
DB 23 AGGATGTCTATGAAATGACGACTCTTCTCTGATACAGCTGAGCTGTACTTGAGC 82
QY 58 tgtgaattctgtggaaagctcgtgtgtgcccctgtgacacatgacatgacatgacatgac 117
DB 83 TTTGGAGATTGTGGAAGAAGGCTGTGTGCGCCACAGAAATTCAGCATTCGATGAATATA 142
QY 118 aagatcattctagaagacatgagagagagagagagagagagagagagagagagagagag 177
DB 143 AAGACAAATCTGTGATGCTTGTCCAGAGAGGTCATGAGGTGACTGACTGACATATTA 202
QY 178 aagccttcgttaattgactacagaagacccctgcgacttgaattgaattgagtgccata 237
DB 203 ACTTCATTTCTTCTGATCCCAACACCCATCTCCCTTTAAATTTGAAATTTGTCTTACA 262
QY 238 ccacagagacagacagagaataatgtttgtacacatgacatgacatgacatgacatgac 291

```

```

DB 263 TCCTTAACCGAAGCTGATTTCAGGATAGCGTCACGCCAACCTGGTTAAGAGATGTCAGAC 322
QY 292 ttgcagagcttaccacccctggcaatcagttatataaattatattttttgttgaata 351
DB 323 ATTGGAAGACACATTTTGGCCACATTTTTCACGCTACAGAAATGATGTGCACATAT 382
QY 352 agagaactttaaaatgatlgtgtgagagcttatactacacacagacatgaagaag 411
DB 383 GGTGACATGATTTAGAAAGTCTCTGTAAGATGCTGTTCTATTAAGAACATTAGAGAAA 442
QY 412 ctacaggaacacacacatgacatgtacagcttataagaccctgttgcctggagagac 471
DB 443 CTACAGGACGACCAAGATTTGATGTTGTTGCTGCGGATGCTATTCGCCCTGTGAGCTA 502
QY 472 atgcctgagttgttccagcgccttctgtgtacacacttgaacactctcgaaggcaat 531
DB 503 CTGCGTAGCTACTTAAATAACCTTGTGTATACAGCTCCATTTCTCTCTGCTGACGA 562
QY 532 atgagcgaagcgtgtggaacacttcacagctccacttccctgtacactgtgcctagaca 591
DB 563 CTGGAAGAAGCATGTGAGAGGTTTCTGTTCCCTCCATTCCTGTTACTGTATATGTCA 622
QY 592 ggaactacagacagaatgacacttcttggaagaagtaaaaatcctaagtcttcaattg 651
DB 623 GAATTAAGAGATCAATAGACTTTCATGAGAGAGTACCAAAATGATCTATGATGTTAT 682
QY 652 ttccacttctgagatcagaattacagatcatcatttttggaagaagtattatagtagca 711
DB 683 TTTGACTTTTGTTCAGAGATGATGTAAGAACTGGGATTCAGTTTACAGTAAAGTT 742
QY 712 ttat 715
DB 743 CTAG 746
RESULT 13
HSAJ5162
LOCUS Homo sapiens mRNA for UDP-glucuronosyltransferase.
DEFINITION AJ005162
ACCESSION AJ005162.1 GI:3135024
VERSION 1
KEYWORDS UDP-glucuronosyltransferase; UGT2B4 gene.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2079)
AUTHORS Rittler, J.K.
TITLE Direct Submission
JOURNAL Submitted (08-MAY-1998) Rittler, J.K., Pharmacology and Toxicology,
of Virginia Commonwealth University, P.O. Box 980613, Medical College
of Virginia, Richmond, Virginia, 23113-0613, USA
REFERENCE 2 (bases 1 to 2079)
AUTHORS Rittler, J.K., Chen, F., Sheen, Y.Y., Lubets, R.A. and Owens, I.S.
TITLE Two human liver cDNAs encode UDP-glucuronosyltransferases with 2
log differences in activity toward parallel substrates including
hydroxycholeic acid and certain estrogen derivatives
JOURNAL Biochemistry 31 (13), 3409-3414 (1992)
MEDLINE 92207964
REFERENCE 3 (bases 1 to 2079)
AUTHORS Jackson, M.R., McCarthy, L.R., Harding, B., Wilson, S., Coughtrie, M.W.
and Burchell, B.
TITLE Cloning of a human liver microsomal UDP-glucuronosyltransferase
JOURNAL CDNA Biochem. J. 242 (2), 581-588 (1987)
MEDLINE 87241362
FEATURES
Source Location/Qualifiers
1..2079
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cclone="UDPgtH-1"
/clone_jib="Lambda bacteriophage gt11"
/tissue_jib="Liver"

```


